

# GIS Modeling of Air Toxics Releases from TRI- Reporting and Non-TRI-Reporting Facilities: Impacts for Environmental Justice

**Dana C. Dolinoy and Marie Lynn Miranda**

**doi:10.1289/ehp.7066 (available at <http://dx.doi.org/>)**

**Online 13 September 2004**



**The National Institute of Environmental Health Sciences  
National Institutes of Health  
U.S. Department of Health and Human Services**

# **GIS Modeling of Air Toxics Releases from TRI-Reporting and Non-TRI-Reporting Facilities: Impacts for Environmental Justice**

Dana C. Dolinoy, MS and Marie Lynn Miranda\*, PhD

Dana C. Dolinoy, MS  
Integrated Toxicology Program and  
Children's Environmental Health Initiative  
Nicholas School of the Environment and Earth Sciences  
Duke University

\* Corresponding Author:

Marie Lynn Miranda, PhD  
Director, Children's Environmental Health Initiative  
Nicholas School of the Environment and Earth Sciences  
A134-LSRC  
Box 90328, Duke University  
Durham, NC 27708  
Phone: 919-613-8023  
FAX: 919-684-8741  
E mail: mmiranda@duke.edu

# GIS Modeling of Air Toxics Releases

## **Key words:**

Geographic Information Systems (GIS), Environmental Justice, Air Dispersion Modeling, Air Toxics, Toxics Release Inventory (TRI)

## **Acknowledgements:**

This research was supported in part by funding from the National Institute of Environmental Health Sciences (ES10356). The authors do not have any competing conflicts of interest. The authors would like to thank Christine Bradshaw, Jay Hamilton, Jonathan Levy, and M. Alicia Overstreet for their help and guidance on this project.

## **Abbreviations:**

CDF - Cumulative Distribution Function

EPA - U.S. Environmental Protection Agency

EPCRA - Emergency Planning and Community Right-to-Know Act

GIS - Geographic Information Systems

$\text{g}/(\text{s}\cdot\text{m}^2)$  – grams per second meter squared

$\mu\text{g}/\text{m}^3$  - micrograms per cubic meter

$\text{ng}/\text{m}^3$  - nanograms per cubic meter

NCDC - National Climatic Data Center

$\text{km}^2$  - square kilometers

SIC - Standard Industrial Classification

TRI - Toxics Release Inventory

UTM -Universal Transverse Mercator

ZCTAs - Zip Code Tabulation Areas

# GIS Modeling of Air Toxics Releases

Abstract

Background

Materials and Methods

Study Area

Demographic Data

Facility Data

Selecting a Base Case

Emission Estimation Algorithm

ISC Dispersion Modeling

Input Requirements

Models

Statistical and Spatial Analysis

Cumulative Distribution Functions

Multivariate Statistics

Kriging

Results

Importance of Inclusive Modeling

Importance of Geographic Resolution

Discussion and Conclusion

References

# GIS Modeling of Air Toxics Releases

## Abstract

The Toxics Release Inventory (TRI) requires facilities with 10 or more full time employees that process over 25,000 pounds in aggregate or use more than 10,000 pounds of any one TRI chemical to report releases annually. However, little is known about releases from non-TRI-reporting facilities; nor has attention been given to the very localized equity impacts associated with air toxics releases. Using Geographic Information Systems (GIS) and Industrial Source Complex dispersion modeling, we developed methods for characterizing air releases from TRI-reporting as well as non-TRI-reporting facilities at four levels of geographic resolution. We characterized the spatial distribution and concentration of air releases from one representative industry in Durham County, NC. Inclusive modeling of all facilities rather than modeling of TRI sites alone significantly alters the magnitude and spatial distribution of modeled air concentrations. Modeling exposure receptors at more refined levels of geographic resolution reveals localized, neighborhood-level exposure hotspots that are not apparent at coarser geographic scales. Multivariate analysis indicates that inclusive facility modeling at fine levels of geographic resolution reveals income and race exposure disparities. These new methods significantly enhance the ability to model air toxics and perform equity analysis and clarify conflicts in the literature regarding environmental justice findings. This work has substantial implications for how to structure TRI reporting requirements, as well as how and what types of analysis will successfully elucidate the spatial distribution of exposure potentials across geographic, income, and racial lines.

# GIS Modeling of Air Toxics Releases

## 1 Background

Temporal, spatial, and other special circumstances can result in disproportionate environmental exposures and may play a role in differential health outcomes. For example, numerous studies highlight the disparate impact of asthma and allergies on specific subsets of the population, including minorities and poor families, as well as people living in urban environments (Oliveti et al. 1996; Weitzman et al. 1990; Wissow et al. 1988). In addition, people of color are more likely to live in areas that fail to meet national ambient air quality standards. In 1990, 57 percent of whites, 65 percent of African Americans, and 80 percent of Hispanics lived in counties that exceeded one of the federal criteria air pollutant standards (Wernette and Nieves 1992). In addition, 12 percent of whites, 20 percent of African Americans, and 31 percent of Hispanics lived in counties that failed to meet the federal standard for three or more criteria air pollutants (Wernette and Nieves 1992).

The Toxics Release Inventory (TRI), created by the Emergency Planning and Community Right-to-Know Act (EPCRA) of 1986, requires manufacturing facilities with 10 or more full time employees that process over 25,000 pounds in aggregate or use more than 10,000 pounds of any one of the 650 TRI chemicals to report their releases and waste management strategies annually to the U.S. Environmental Protection Agency (EPA) (U.S. Environmental Protection Agency 2001). Reporting requirements cover emissions from routine processing and/or accidental releases as well as chemicals managed as waste from businesses categorized in Standard Industrial Classification (SIC) codes 10, 12, 20-39, 49, and 51 including metal and coal mining, printing, chemical, paper, electronic, plastics, refining, metal, and other industries. Facilities that fail to report annual releases by July 1 each year are subject to fines of up to \$27,500 per day (Office of Environmental Information 2001).

## GIS Modeling of Air Toxics Releases

The purpose of EPCRA is twofold. First, it seeks to provide citizens with information on chemical releases and waste management activities in and around their neighborhood, thereby empowering them to hold companies accountable for their emissions (Office of Environmental Information 2001). Second, it attempts to provide government agencies with data for research and policy development (Office of Environmental Information 2001). While an important advance in helping communities to understand the local air toxics load, the TRI program suffers from at least three weaknesses. First minimum reporting requirements do not require smaller industrial facilities to report. Theoretically, cumulative effects of smaller non-TRI-reporting facilities might outweigh the individual effect of larger (but fewer) TRI-reporting facilities. Second, EPA's TRI database (as well as TRI data organized and maintained by environmental interest groups) does not address environmental fate and transport of industry emissions using modeling and other analytical techniques. The characteristics of pollutant concentration distributions depend on a variety of factors including media emitted, physical properties of the chemical, wind direction and speed, meteorological conditions, and stack height. Finally, by reporting emissions at the county level, the TRI database fails to capture important highly localized equity impacts.

The analysis presented here develops a methodology to characterize releases inclusively by incorporating emissions from smaller, non-TRI-reporting facilities. In addition, the project is carried out at four geographic levels of resolution (zip code, Census tract, Census block group, and Census block) in order to assess the importance of geographic resolution in analyzing air toxics emissions. The paper models pollutant concentrations using dispersion modeling and geographic information systems (GIS) analysis and hypothesizes that *inclusive modeling of releases from TRI-reporting and smaller non-TRI-reporting facilities at finer levels of*

# GIS Modeling of Air Toxics Releases

*geographic resolution will change the distribution of exposure potential for people living in a given area and subsequently improve the quality of equity analysis.*

## **2 Materials and Methods**

### **2.1 Study Area**

The analysis focuses on Durham County, North Carolina, which represents a broad range of values across demographic, socioeconomic, and social indicators. Table 1 shows selected demographics for Durham County from 2000 Census data. Forty percent of Durham County residents are black and almost eight percent are Hispanic (U.S. Census Bureau 2003). In addition, approximately ten percent of residents are living in poverty and almost forty-six percent reside in rental properties (U.S. Census Bureau 2003). Compared to Durham County as a whole, the State of North Carolina, and the United States, Central Durham has a higher percent of minorities, higher percent of families in poverty, higher percent of children under the age of six in poverty, lower median household income, and a higher percent of renter occupied housing (U.S. Census Bureau 2003). Figure 1 highlights the location of Durham County within North Carolina. The yellow box represents Central Durham. Figure 2 depicts two demographic variables for Durham County (percent African American by 2000 Census block and median household income by 2000 Census block group). Note that areas within the yellow box (Central Durham) are characterized with a higher percentage of African Americans and lower household median income.

### **2.2 Demographic data**

This project uses year 2000 demographic data from the U.S. Census Bureau. Census demographic information is available in four different geographic scales: zip code tabulation areas (ZCTAs), tracts, block groups, and blocks. Zip codes and tracts designate the largest



## GIS Modeling of Air Toxics Releases

geographic areas. The most detailed and focused information is contained in blocks. Blocks are also combined into block groups, an intermediate category. Durham County contains 20 zip codes, 53 tracts, 129 Census block groups, and 3284 Census blocks. Figure 3 depicts zip code, tract, block group, and block boundaries for Durham County.

### **2.3 Facility Data**

Year 2000 TRI data were extracted from EPA's TRI Explorer and uploaded into a GIS. TRI facility locations and emissions were geocoded to a base map using latitude and longitude. Facility location was cross-referenced against tax parcel data to ensure accurate geolocation. TRI data from 2000 indicate that North Carolina is home to 874 TRI sites, releasing over 126 million pounds of contaminants to the air. Durham County is home to 16 of these sites releasing over 48,000 pounds of contaminants to the air.

Non-reporting facilities within TRI-reporting SIC codes were extracted from city marketing directories and imported into the GIS project. Facility locations were address-geocoded to the individual tax parcel unit. City marketing data contain countywide listings of business names, addresses, contact persons, employee range, and SIC codes. Year 2000 city marketing data indicate that Durham County contains over 400 non-TRI-reporting industrial facilities classified in TRI-SIC reporting codes. Figure 4 maps facilities required to report to TRI and those in the same SIC codes, which were not required to report to TRI in Durham County in 2000. Note that while only three of the TRI-reporting facilities are located in the low-income, predominantly minority communities of Central Durham, the majority of the non-TRI-reporting, smaller facilities are situated in Central Durham.

## GIS Modeling of Air Toxics Releases

### 2.4 Selecting a Base Case

Ideally, all SIC codes and pollutants subject to TRI guidelines would be evaluated in an aggregate model. However, to facilitate the development of the spatial methods described here, we selected a prototype SIC code and pollutant to model. Based on an evaluation of year 2000 air releases from several Southern states, a 4-digit rather than 2-digit SIC code was selected as a base case. This selection criterion better supported the required modeling assumptions that a TRI facility within a defined SIC code releases a similar pollutant profile when compared to other TRI and non-TRI reporting facilities in the same SIC code.

Of the 16 TRI-reporting sites in Durham County, no two facilities were defined within one 4-digit SIC code. Therefore, we based the prototype selection on quantity and type of air releases. SIC code 2752, representing Commercial Lithographic Printing, was chosen because it released one type of pollutant rather than multiple classes of pollutants. In addition, it ranked second within Durham County for total air releases. Within Durham County, SIC code 2752 contained 1 TRI-reporting and 36 non-TRI-reporting sites. The TRI-reporting site was located in Southern Durham, while the 36 non-TRI code 2752 sites were spread throughout the county – see Figure 5. The TRI-reporting site emitted “certain glycol ethers” as fugitive (rather than stack) releases to the air.

An evaluation of air emissions in six Southern states with manufacturing facilities classified in SIC code 27 revealed wide variability among 2-digit SIC codes. However, variability was low when facilities were restricted to 4-digit SIC codes. For example, of the 16 Virginian facilities in 2-digit SIC code 27 releasing TRI chemicals to the air in 2000, three were identified within 4-digit SIC code 2752. All three facilities released “certain glycol ethers” or

## GIS Modeling of Air Toxics Releases

ethylene glycol to the air. In addition, for all three facilities, the overwhelming majority of the releases were identified as fugitive rather than stack.

Glycol ethers represent a bundle of chemicals used in many industries (including printing) as solvents (Environmental Defense 2002). According to 2000 TRI data, the TRI-reporting facility (PBM Graphics) in Durham County released 13,733 pounds of “certain glycol ethers” to the air. All releases were classified as fugitive, rather than point source. In order to determine the specific chemical used, PBM Graphics was contacted directly. According to PBM Graphics personnel, the company used ethylene glycol monobutyl ether in 2000. Ethylene glycol monobutyl ether is a non-photo-reactive volatile organic compound used as a solvent during printing processes. It is a suspected cardiovascular, blood, developmental, endocrine, gastrointestinal, kidney, neurological, and respiratory toxicant (Environmental Defense 2002; New Jersey Department of Health and Senior Services 2001). According to the CalTox multi-media and multi-exposure model, ethylene glycol monobutyl ether exposure potential is primarily through air (rather than settled materials) and induces health effects via inhalation and ingestion (State of California Department of Toxic Substances Control 2002). The half-life of ethylene glycol monobutyl ether is approximately 16 hours in the air. Both the exposure properties and half-life make ethylene glycol monobutyl ether an appropriate pollutant for air dispersion modeling. Although other facilities within SIC code 2752 may emit different subtypes of “certain glycol ethers,” the exposure properties and half lives of individual compounds do not vary significantly enough to invalidate modeled concentrations.

### **2.5 Emissions Estimation Algorithm**

Comparing TRI-reporting to non-TRI-reporting sites requires the estimation of emissions from non-TRI-reporting facilities. Ideally, annual averages of production units would be utilized.

## GIS Modeling of Air Toxics Releases

However, employees can serve as a proxy when production units are not readily known.

Therefore, we generated an employee-based emissions algorithm in order to impute emissions to non-TRI-reporting facilities. The general algorithm for computing the estimated emissions is as follows. Step 1 involves the calculation of a per-employee emissions rate for a chemical of concern ( $C_i$ ) based on data from all TRI-reporting facilities ( $F_1, F_2, F_3 \dots F_n$ ). In Step 2 emissions are imputed to non-TRI-reporting facilities by multiplying the per employee emissions rate (based on the TRI-reporting facilities) by the number of employees working at each of the non-TRI-reporting facilities.

Using employees instead of production units fails to address how economies of scale might affect production patterns. However, the literature (Dasgupta et al. 2002; Little et al. 1987; U.S. Environmental Protection Agency 2001) indicates that smaller facilities tend to emit more pollution on a per unit of production basis than larger units, so our imputed data likely represent an underestimate of emissions from non-TRI-reporting facilities. Therefore, if significant differences are noted between TRI-reporting models and inclusive TRI-reporting plus non-TRI-reporting sites, under this conservative approach, the actual effect can be assumed to be greater.

In order to impute emissions for the 36 non-TRI-reporting facilities in code 2752, a per employee emissions rate (based on emissions from PBM Graphics) was calculated. PBM graphics employed approximately 375 employees in its printing facility in 2000. The reported emissions were 13,733 pounds of glycol ethers delivering an imputed emissions rate of 36.6 pounds per employee. As a data check, several SIC code 2752 facilities in surrounding states had similar emissions per employee ratios. City marketing directories were used to ascertain the number of employees at the 36 non-TRI-reporting facilities. Using the per employee emissions rate and the number of employees, we imputed air emissions for the 36 non-TRI-reporting

## GIS Modeling of Air Toxics Releases

facilities. Once the base-case methodology is developed, future analysis should include Monte Carlo simulations that vary the per employee emissions rate across non-TRI-reporting facilities. Figure 5 depicts the 37 facilities in Durham County classified in SIC code 2752 and their corresponding imputed emissions. According to imputed emissions estimates, the 36 non-TRI-reporting facilities emitted a total of 22,156 pounds per year.

### **2.6 ISC Dispersion Modeling**

Developed by EPA, the Industrial Source Complex (ISC) model is one of the most widely used and successful steady state Gaussian-based air dispersion models. Major assumptions of Gaussian models include: 1) that the rate of plume diffusion is proportional to contaminant concentration; 2) a constant emissions rate; 3) a conservative pollutant (no chemical reactions or biodecay); 4) relatively flat terrain; and 5) perfect ground reflection. Because of these assumptions, Gaussian models are most appropriate for local applications within 50 kilometers or 2500 square kilometers ( $\text{km}^2$ ) (Masters 1998). Gaussian models incorporate two dispersion coefficients based on the standard deviations of the horizontal and vertical Gaussian distributions of the downwind plume dispersion (Masters 1998). The standard deviations or dispersion coefficients increase with distance downwind of the source. In addition to distance, the dispersion coefficients also consider atmospheric stability parameters that address qualitative descriptions of prevailing weather conditions such as season, time of day, and degree of cloud cover. Gaussian model output reports annual average concentration of pollutant for each defined receptor. Receptors are user defined areas of interest and often include the geographic centroid of zip codes or Census tracts or points on regular grids spanning a study region.

This analysis utilized the short term ISC model, ISCST3, which captures initial mixing phenomena at the source, and is best suited for study areas less than  $2500 \text{ km}^2$ . Containing

## GIS Modeling of Air Toxics Releases

approximately 751 km<sup>2</sup>, Durham County falls well within this limitation. ISCST3 short-term modeling allows for multiple source and receptor specifications and requires users to input a year's worth of hourly meteorological data from the National Weather Service. The modeling period was January 1, 2000 to December 31, 2000. Source emission rates were treated as constant over one year by converting the number of pounds released in calendar year 2000 to grams per second.

### *2.6.1 Input Requirements*

Within ISCST3, sources may be specified as POINT (stack), AREA (storage piles or irregular shapes), or VOLUME (multiple vents or conveyor belts) types (Office of Air Quality Planning and Standards 1995). According to TRI Explorer and conversations with TRI and non-TRI-reporting sites in SIC code 2752, releases from printing lithography facilities are generally fugitive rather than point. The AREA rather than VOLUME type was selected in order to control for differences in building size. The area of each building footprint (in square meters) was based on the area of PBM graphics scaled to the number of employees. Furthermore, the AREA subtype is best suited for low level releases with no plume rise (Office of Air Quality Planning and Standards 1995). Area emission rates based on the per employee emissions algorithm were entered into ISCST3 in grams per second meter squared (g/(s-m<sup>2</sup>

A year's worth of hourly meteorological data was compiled using the PCRAMMET meteorological preprocessor program. Ground level weather data was obtained from the Raleigh-Durham Airport surface station (approximately 3 miles southwest of Durham County). The anemometer height at the airport is approximately 10.1 meters. Mixing height data from the

## GIS Modeling of Air Toxics Releases

National Climatic Data Center (NCDC) was obtained for the closest station, Greensboro, North Carolina, located approximately 50 miles west of Durham County. The twice daily mixing height values were combined with hourly surface data using PCRAMMET in order to derive hourly interpolated mixing height values.

Receptors were defined as the geographic centroid of four modeling units (zip code, Census tract, Census block group, and Census block) in Durham County, North Carolina. For each geographic scale, receptor coordinates of the centroids were entered as UTM coordinates. The default receptor elevation of ground level was utilized. Areas adjacent to Durham County were not specified or analyzed in this study.

In order to control for pollutant fate and transport characteristics, the ISCST3 models allows users to input pollutant half-life and specify average land terrain across the study area of interest. As an initial exercise, default values for pollutant half-life and landscape terrain were first specified. The ISCST3 default pollutant half-life is 4 hours. However, the half-life of ethylene glycol monobutyl ether is 16 hours. When the default half-life was changed to 16 hours, significant changes in output results were not observed. The default ISCT3 landscape terrain is rural. Because the landscape of Durham County is variable, the models were also specified using an urban terrain. Results did not vary significantly based on rural versus urban specified landscape terrain.

### 2.6.2 *Models*

In order to assess the importance of inclusive modeling of all emitters as well as the importance of geographic resolution, eight ISCT3 models were specified:

- 1a: TRI-reporting sites alone at zip code level (20 receptors)
- 1b: All emitters at zip code level (20 receptors)

## GIS Modeling of Air Toxics Releases

- 2a: TRI-reporting sites alone at tract level (53 receptors)
- 2b: All emitters at tract level (53 receptors)
- 3a: TRI-reporting sites alone at block group level (129 receptors)
- 3b: All emitters at block group level (129 receptors)
- 4a: TRI-reporting sites alone at block level (3824 receptors)
- 4b: All emitters at block level (3824 receptors)

Each model run generated an average concentration of the pollutant in micrograms per cubic meter ( $\mu\text{g}/\text{m}^3$ ) in a particular receptor grid for the entire year. Comparing the (a) to (b) runs (e.g. 1a to 1b) allows for determination of the importance of inclusive modeling, i.e. does including non-TRI-reporting facilities change exposure potential across different demographic groups. Comparing the (1)-(4) runs explores the importance of geographic resolution in analyzing contaminant distribution across demographic groups. Statistical and spatial analysis of pollutant concentration was based on model output and was not verified by collecting environmental samples.

### **2.7 Statistical and Spatial Analysis**

Using spatial and tabular tools within GIS, modeled emissions from the ISC dispersion models were aggregated into spatially referenced datasets and combined with underlying Census demographic data. The combined datasets were imported into Microsoft Excel and STATA for statistical analysis.

#### *2.7.1 Cumulative Distribution Functions*

Comparing means (or medians) of modeled concentration, as is the case in multivariate regression analysis, may overemphasize differences near the center of the concentration



## GIS Modeling of Air Toxics Releases

distribution. Many times, scientists and public health analysts are most concerned with areas of relatively high (or low) exposure risk (i.e., the tails of the distribution). Therefore, we utilized cumulative distribution functions (CDF) to address disparate exposure potential across two groups. For illustrative purposes, the y-axis of a CDF curve represents the percentage of the population (from 0 to 100) and the x-axis represents exposure potential (interpolated monitoring site data or modeled concentration levels for example) (Lopez 2002).

The CDFs were estimated using techniques reported in Waller et al. 1999 (Waller 1999). CDFs were created for all four levels of geographic resolution (zip code, Census tract, Census block group, and Census block), although only data for Census block, the finest level of resolution, are presented here. Population data for each subpopulation of interest for each geographic level of resolution were determined using 2000 Census data and GIS. Exposure values were assigned to each unit based on modeled concentration values for the corresponding geographic level of resolution.

### 2.7.2 *Multivariate Statistical Analysis*

In order to access multiple demographic variables at once, multivariate statistical analysis was performed with STATA 8.0. All dependent variables were log-transformed to normalize right skewed data.

### 2.7.3 *Kriging*

Using the spatial analyst extension within ArcView 3.2 GIS software, a set of contour lines representing predicted concentration of glycol ethers for the entire County of Durham were developed. Using a built-in geostatistical program with user-defined parameters, ArcView software interpolates lines that represent locations with the same pollutant concentration

## GIS Modeling of Air Toxics Releases

magnitude. Although kriging the modeled data introduces an additional layer of uncertainty, the smoothed contour lines depict a more easily interpreted array of pollutant concentrations, which is extremely useful for neighborhood level equity analyses and community outreach.

### 3 Results

The outcome variables of interest from the ISC dispersion modeling datasets include: 1) the annual average concentration ( $\mu\text{g}/\text{m}^3$ ) of glycol ethers based on modeling of the TRI site alone; and 2) the annual average concentration ( $\mu\text{g}/\text{m}^3$ ) of glycol ethers based on modeling of the TRI site plus non-TRI-reporting sites. Modeling was conducted at four geographic levels: zip code, Census tract, Census block group, and Census block. The annual average concentrations were converted to nanograms per cubic meter ( $\text{ng}/\text{m}^3$ ) in order to facilitate presentation and log transformation.

Table 2 presents descriptive statistics for the annual average concentration in zip codes, Census tracts, Census block groups, and Census blocks. Dispersion modeling data range in concentration from  $0.3 \text{ ng}/\text{m}^3$  (TRI site alone, zip code level) to  $821.05 \text{ ng}/\text{m}^3$  (all sites together, Census block level). The inclusion of non-TRI-reporting sites in the model increases the average concentration among zip codes from  $1.5 \text{ ng}/\text{m}^3$  to  $4.9 \text{ ng}/\text{m}^3$ . Likewise, the inclusion of smaller non-TRI-reporting sites in the model increases the average concentration among Census tracts from  $2.7 \text{ ng}/\text{m}^3$  to  $10.3 \text{ ng}/\text{m}^3$ ; for Census block groups from  $3.0 \text{ ng}/\text{m}^3$  to  $10.2 \text{ ng}/\text{m}^3$ ; and Census blocks from  $3.9 \text{ ng}/\text{m}^3$  to  $12.1 \text{ ng}/\text{m}^3$ .

#### 3.1 Importance of Inclusive Modeling

As shown in Figure 6, inclusive modeling of all facilities, accomplished by imputing emissions to non-TRI-reporting facilities in the same SIC code, rather than modeling of TRI sites alone, significantly alters the magnitude and spatial distribution of modeled air concentrations.

## GIS Modeling of Air Toxics Releases

Recall from Figure 2 that areas in southern Durham County have higher household median incomes and relatively low densities of minorities as compared to Central Durham, as measured by Census 2000 data. Note the northward drift of higher concentration contours (the deeper the red color, the higher the modeled concentration) into lower income, predominantly minority communities. Thus incorporating the non-TRI-reporting facilities provides a substantially different perspective on exposure to contaminants across race and income lines.

For non-inclusive modeling at each level of geographic resolution, major impacts occur within a few miles of the TRI site. For inclusive modeling of the TRI-reporting plus non-TRI-reporting facilities at each level of geographic resolution, major impacts are spread throughout Durham County and into adjacent counties (Chatham, Orange, and Wake). The aggregate effects of modeling multiple smaller non-TRI-reporting emissions in Central Durham are of the same order of magnitude as the effects of the larger TRI site in southern Durham. Although non-TRI-reporting sites do not significantly affect exposure potential in areas with TRI facilities, they do affect exposure potential in areas at some distance from TRI facilities. This results in part from the size and specific locations of the reporting and non-TRI-reporting 2752 facilities and may not necessarily hold when generalized to other SIC codes.

Figure 7 depicts the CDF for black and white subpopulations modeled at the block level. Graph A represents exposure values for non-inclusive modeling of air emissions for TRI-reporting facilities only. The CDF curves for black and white subpopulations exhibit a narrow gap, indicating that a slightly larger proportion of whites reside in blocks with lower exposure potentials. Graph B represents exposure potential values for inclusive modeling of air emissions from all emitters at the block level. Inclusive modeling results in a larger gap between the CDF curves for the black and white subpopulations. The increase in exposure disparity as you move

## GIS Modeling of Air Toxics Releases

from non-inclusive to inclusive modeling persists at the three coarser levels of geographic resolution (data not shown). However, the increased gap is most apparent at the Census block level.

Figure 8 depicts the CDF curves for comparing adult and non-adult (under 18) subpopulations. Again, graph A represents exposure values for non-inclusive modeling of air emissions for TRI-reporting facilities only. The CDF curves for adults and non-adults overlap, indicating a lack of disparate exposure. Graph B represents exposure values for inclusive modeling of air emissions from all emitters. Unlike the CDF for race depicted above, potential exposure disparities based on age do not appear sensitive to non-inclusive versus inclusive modeling. CDFs were also estimated for persons under five versus persons over five. Significant differences in exposure potential were not observed based on non-inclusive versus inclusive modeling.

### **3.2 Importance of Geographic Resolution**

Comparing run 1 (zip code receptor) through run 4 (Census block receptors) in Table 2 reveals that modeling receptors at a more refined geographic resolution alters the annual average concentrations of glycol ethers for both the TRI models alone as well as the inclusive all sites models. The range of concentrations for the Census block model is successively greater than the range of concentrations for the coarser geographic scale models. As the same emissions are being spread over successively smaller areas, the wider range of concentrations at finer geographic scales is an intuitive result.

Figure 9 depicts kriging results with two contour maps representing annual average concentration of glycol ethers for Durham County, North Carolina (the deeper the blue color, the higher the modeled concentration). Both maps represent inclusive modeling of all emitters.

## GIS Modeling of Air Toxics Releases

However, the left figure depicts contours based on modeling sources and receptors at the coarser zip code level of geographic resolution, while the right figure represents modeling at the finer Census block group level. As shown in Figure 9, modeling exposure receptors at finer geographic levels of resolution (i.e. Census block group rather than zip code) reveals localized, neighborhood-level exposure hotspots that are not apparent at coarser geographic scales – note in particular the high concentration contours that appear in Central Durham under this alternative modeling approach. Modeling finer geographic levels of resolution provides a substantially different perspective on exposure to contaminants across race and income lines. Unlike the zip code level contours, the Census block group level contours highlight areas in Central Durham, characterized by a higher percentage of minorities and a lower median household income, as potential hotspots for exposure.

To better summarize and assess if modeling of air emissions at varying levels of geographic resolution affects the distribution of exposure potential, we performed multivariate statistical analysis on the relationship between concentration and race and income. We focus specifically on these two variables because of their ubiquitous use in equity analysis. Dependent variables of interest were the modeled concentrations at the four geographic levels of resolution (zip code, Census tract, Census block group, and Census block).

Tables 3 and 4 summarize regression results across geographic scale. In Table 3, moving from top to bottom indicates stepping from coarser (zip code) to finer (Census block) levels of geographic resolution. A positive sign indicates a positive coefficient on the regression coefficient and an “S” indicates significance at the 0.05 level. Likewise, a negative sign indicates a negative coefficient on the regression coefficient and a “NS” indicates lack of significance at the 0.05 level. As you move from top to bottom, both the income and minority

## GIS Modeling of Air Toxics Releases

variables become significant and of the expected sign. These results highlight the importance of spatial resolution in conducting equity analysis. Table 4 presents more detailed regression results from multivariate statistical analysis. Comparing the zip code to block models explores the importance of geographic resolution in analyzing contaminant distribution. The Census block model, representing inclusive modeling of all emitting sites at a very refined geographic scale, indicates exposure potential disparities across both income and race.

Additional multivariate analyses including the relationships between modeled concentration and percent children, percent vacant housing, and percent receiving public assistance did not reveal any statistically significant trends (results not presented here). In addition, percent minority appears to be the most relevant “race” variable based on the modeled data; and median household income appears to be the most important “income” variable based on the modeled data (results not presented here).

### **4 Discussion and Conclusion**

Previous air toxics and TRI studies have taken advantage of advances in spatial and statistical mapping software to assess how geographic levels of resolution affect environmental justice conclusions. In a 2002 article summarizing many existing geographic-based air toxics studies, Lopez explains that conclusions often differ depending on the geographic unit of analysis (Lopez 2002). For example, “micro area” studies that observe areas with and without facilities conclude that race is not a significant predictor for site but that income may play a role. On the other hand, “meso area” studies that expand the area of interest to include blocks adjacent to facilities often conclude that race is an important predictor for siting and income is not. Furthermore, results from “macro” level studies that compare counties to other counties or states to other states have correlated industrial facility siting with large percentages of minorities and

## GIS Modeling of Air Toxics Releases

persons in poverty. However, the results may be confounded by urban/rural status and other trends. Other traditional environmental justice analysis of industrial siting has focused on the location of facilities and not on concentration distributions and subsequent exposure potential (Morello-Frosch et al. 2002).

Our study attempts to clarify conflicts in the literature regarding facility siting, exposure potential, and equity by developing methods for inclusive modeling of releases at fine levels of geographic resolution. Although the results described here are specific to emissions of glycol ethers from Printing/Lithography sites in Durham County in 2000, the method developed is relevant across time, space, and industries. This is one of the first studies to develop methods for characterizing and mapping releases from smaller, non-TRI-reporting facilities. The study methodology further characterizes pollutant distribution and fate and transport by incorporating atmospheric dispersion modeling. The use of GIS as a platform for data storage, statistical analysis, and kriging remains an important cornerstone for conducting spatially based environmental justice research. In addition, although average and maximum annual average concentrations of the pollutant (Table 2) do not approach the non-Cancer levels of concern set by California EPA (700,000 ng/m<sup>3</sup>) and EPA (20,000 ng/m<sup>3</sup>), the methods developed and presented here represent an innovative prototype for contaminant analysis. A full characterization of exposure potential would take into consideration releases from other sources and adjacent counties.

In a recent article, Maantay attributes the failure of previous studies to address small polluters such as automobile repair shops to the lack of standardized and publicly available datasets on small polluters (Maantay 2002). Although there are several caveats to using employees to estimate facility emissions, we believe that the approach offers a sound and

## GIS Modeling of Air Toxics Releases

creative solution for addressing these data limitations. However, future studies adopting this mechanism should perform some quality assurance techniques to ensure that number of employees reflects a representative proxy to production units and/or pollutant emissions.

The results indicate that the inclusive modeling of all facilities significantly alters the magnitude and spatial distribution of modeled air concentrations (Figure 6). Modeling all sites together rather than modeling TRI sites alone increases the magnitude of modeled concentrations – especially in areas with no TRI facilities. The red concentration contours depicted on the inclusive map are spatially correlated to high minority and low income neighborhoods presented in Figure 2. The same correlation is not observed for the TRI-reporting contours. In addition, the CDF curves indicate that inclusive facility modeling at fine levels of geographic resolution results in exposure disparities across race, but not age (Figures 8 and 9).

As described above, a significant body of literature exists comparing varying levels of geographic resolution to different exposure potential outcomes (Glickman 2004; Lopez 2002; Maantay 2002; Sheppard et al. 1999). Intuitively, it makes sense that the finer the geographic resolution, the higher the predicted exposure concentration. This paper attempts to build upon these studies by showing that the choice of geographic resolution significantly affects both the significance and trend of multivariate statistical analysis of underlying demographics. In our study, concentration gradients are substantially influenced by the resolution of the model, indicating that receptor choice is a significant modeling parameter and that localized equity impacts may be best represented at the block level.

These new methods significantly enhance the ability to model air toxics and perform equity analysis. They also clarify conflicts in the literature regarding environmental justice findings. In modeling air toxics, both the fate and transport literature and the mechanistic



## GIS Modeling of Air Toxics Releases

literature indicate that modeling inclusively at a refined geographic scale makes biological sense. From a policy standpoint, then, it becomes critical to understand how the design of spatial analyses can shape conclusions. This work, as it moves forward, will also have much to say about which facilities should be required to report to TRI, as well as how much reliability we can place on current TRI data. Future analysis will consider multiple contaminant exposures from multiple industries and explore the use of Toxic Equivalency Factors to better analyze underlying justice concerns and exposure potential.

## 5 References

Dasgupta S, Lucas R, Wheeler, D. 2002. Plant size, industrial air pollution, and local income: evidence from Mexico and Brazil. *Environment and Development Economics* 7(2): 365-381.

Environmental Defense. 2002. Glycol Ethers. New York, NY, Environmental Defense. Available: [http://www.scorecard.org/chemical-profiles/html/glycol\\_ETHERS.html](http://www.scorecard.org/chemical-profiles/html/glycol_ETHERS.html) [accessed 4 October 2002].

Glickman, TS. 2004. Evaluating environmental equity in Allegheny County. New York, NY: Program for the Human Environment, The Rockefeller University. Available: [http://phe.rockefeller.edu/comm\\_risk/comm\\_rsk3.html](http://phe.rockefeller.edu/comm_risk/comm_rsk3.html) [accessed 6 July 2004].

Little IMD, Mazumdar D, Page JM. 1987. *Small Manufacturing Enterprises: A Comparative Study of India and Other Economies*. New York, Oxford University Press.

Lopez R. 2002. Segregation and black/white differences in exposure to air toxics in 1990. *Environ Health Perspect* 110(S2): 289-295.

Maantay J. 2002. Mapping environmental injustices: pitfalls and potential of geographic information systems in assessing environmental health and equity. *Environ Health Perspect* 110(S2): 161-171.

Masters G. 1998. *Air Pollution. Introduction to Environmental Engineering and Science*. Englewood Cliffs, NJ, Prentice Hall: 406-426.

Morello-Frosch R, Pastor M, Porras C, Sadd J. 2002. Environmental justice and regional inequity in Southern California: implications for future research. *Environ Health Perspect* 110(Supplement 2): 149-154.

New Jersey Department of Health and Senior Services. 2001. Hazardous Substance Fact Sheet: 2-butoxy ethanol. Trenton, NJ: New Jersey Department of Health and Senior Services. Available: <http://www.state.nj.us/health/eoh/rtkweb/0275.pdf> [accessed 4 October 2002].

Office of Air Quality Planning and Standards. 1995. *User's Guide for the Industrial Source Complex (ISC3) Dispersion Models*. Research Triangle Park, U.S. Environmental Protection Agency.

Office of Environmental Information. 2001. *The Emergency Planning and Community Right-to-Know Act Section 313 Release and Other Waste Management Reporting Requirements*. Washington D.C., U.S. Environmental Protection Agency.

Olivetti J, Kercsmar CM, Redline S. 1996. Pre- and perinatal risk factors for asthma in inner city African-American children. *Am J Epidemiol* 143(6): 570-577.

## GIS Modeling of Air Toxics Releases

Sheppard E, Leitner H, McMaster R, Tian H. 1999. GIS-based measures of environmental equity: exploring their sensitivity and significance. *Journal of Exposure Analysis and Environmental Epidemiology* 9(1): 18-28.

State of California Department of Toxic Substances Control. 2002. CalTox model description. Sacramento, CA: State of California Department of Toxic Substances Control. Available: [http://www.dtsc.ca.gov/ScienceTechnology/ctox\\_model.html](http://www.dtsc.ca.gov/ScienceTechnology/ctox_model.html) [accessed 10 October 2002].

U.S. Census Bureau. 2003. Census Bureau Home Page. Washington, DC: U.S. Census Bureau. Available: <http://www.census.gov/> [accessed 10 April 2003].

U.S. Environmental Protection Agency. 2001. What is the TRI? Washington, DC: U.S. Environmental Protection Agency. Available: <http://www.epa.gov/tri/general.htm> [accessed July 2001].

Waller LA, Louis TA, Carlin BP. 1999. Environmental justice and statistical summaries of differences in exposure distributions. *Journal of Exposure Analysis and Environmental Epidemiology* 9(1): 56-65.

Weitzman M, Gortmaker S, Sobol A. 1990. Racial, social, and environmental risks for childhood asthma. *Am J Dis Child* 144(11): 1189-1194.

Wernette D, Nieves L. 1992. Breathing polluted air. *EPA Journal* 18: 16-17.

Wissow L, Gittelsohn AM, Szklo M, Starfield B, Mussman M. 1988. Poverty, race, and hospitalization for childhood asthma. *Am J Public Health* 78(7): 777-782.

## GIS Modeling of Air Toxics Releases

**Table 1.** Demographics of Central Durham, Durham County, State of North Carolina, and United States

Region	Population	Hispanic (%)	Black (%)	Families in Poverty (%)	Under 6 in Poverty (%)	Median Household Income (%)	Renter Occupied (%)
Central Durham	57,690	12.5	60.5	22.2	37.8	\$36,368	64.8
Durham County	223,314	7.6	39.5	9.8	19.9	\$43,337	45.7
North Carolina	8,049,313	4.7	21.6	9.0	17.8	\$39,184	30.6
United States	281,421,906	12.5	12.3	9.2	18.1	\$41,994	33.8

Source: U.S. Census 2000 (U.S. Census Bureau, 2003 #491).

## GIS Modeling of Air Toxics Releases

**Table 2.** Year 2000 annual average concentration of ethylene monobutyl ether (nanograms/m<sup>3</sup>)

Geographic Resolution	Number of Receptors	Model Run	Mean	Median	Max.	Min.	SD
Zip Code	20	1a: TRI Alone	1.5	0.9	5.4	0.3	1.31
		1b: All Emitters	4.9	2.8	17.9	0.8	4.4
Census Tract	53	2a: TRI Alone	2.7	2.1	19.8	0.6	2.8
		2b: All Emitters	10.3	9.0	46.4	2.2	7.5
Census Block Group	129	3a: TRI Alone	3.0	2.2	28.2	0.5	3.3
		3b: All Emitters	10.2	9.1	49.4	1.6	7.0
Census Block	3824	4a: TRI Alone	3.9	2.2	799.2	0.4	19.6
		4b: All Emitters	12.1	8.7	821.1	1.2	27.3

## GIS Modeling of Air Toxics Releases

**Table 3.** Trends from the multivariate statistical analysis

	Inclusive Modeling
Zip Code	Minority: +/S Income: +/S
Census tract	Minority: +/NS Income: -/~S
Census block group	Minority: +/S Income: -/NS
Census block	Minority: +/S Income: -/S

S: significance at the 0.05 level  
NS: lack of significance at 0.05 level  
+: positive sign on beta coefficient, positively proportional to pollutant concentration  
-: negative sign on beta coefficient, inversely proportional to pollutant concentration

GIS Modeling of Air Toxics Releases

Table 4. Inclusive modeling regression results of at four levels of geographic resolution from multivariate statistical analysis

	Zip Code Coefficient (p-value)	Tract Coefficient (p-value)	Block Group Coefficient (p-value)	Block Coefficient (p-value)
Constant	-1.018	2.306	2.066	2.0245
Percent Minority	0.0382 <b>(0.0001)</b>	0.360 (0.212)	0.519 <b>(0.011)</b>	0.622 <b>(0.0001)</b>
Household Median Income	0.0000193 (0.022)	-7.89e-06 (0.081)	-4.03e-06 (0.153)	-3.55e-06 <b>(0.0001)</b>
Adjusted R-squared	0.50	0.15	0.13	0.11

## GIS Modeling of Air Toxics Releases

**Figure 1.** Durham County, North Carolina

**Figure 2.** Two demographic variables for Durham County: Percent African American and median household income

**Figure 3.** Boundaries for zip codes, Census tracts, Census block groups, and Census blocks for Durham County

**Figure 4.** TRI-reporting and non-TRI-reporting facilities in all TRI SIC codes in Durham County

**Figure 5.** TRI-reporting and non-TRI-reporting facilities in SIC Code 2752 (Printing-Lithography) and their estimated emissions

**Figure 6.** Modeled air emissions of certain glycol ethers for TRI-reporting (left) and non-TRI-reporting (right) facilities

**Figure 7.** Cumulative distribution function (CDF) curves of modeled Census block level exposure for black and white subpopulations for TRI Sites Alone (A) and All Emitters (B)

**Figure 8.** Cumulative distribution function (CDF) curves of modeled Census block level exposure for child and adult subpopulations for TRI Sites Alone (A) and All Emitters (B)

**Figure 9.** Modeled inclusive air emissions of certain glycol ethers for all facilities at the Census tract (left) and Census block group (right) level



## GIS Modeling of Air Toxics Releases

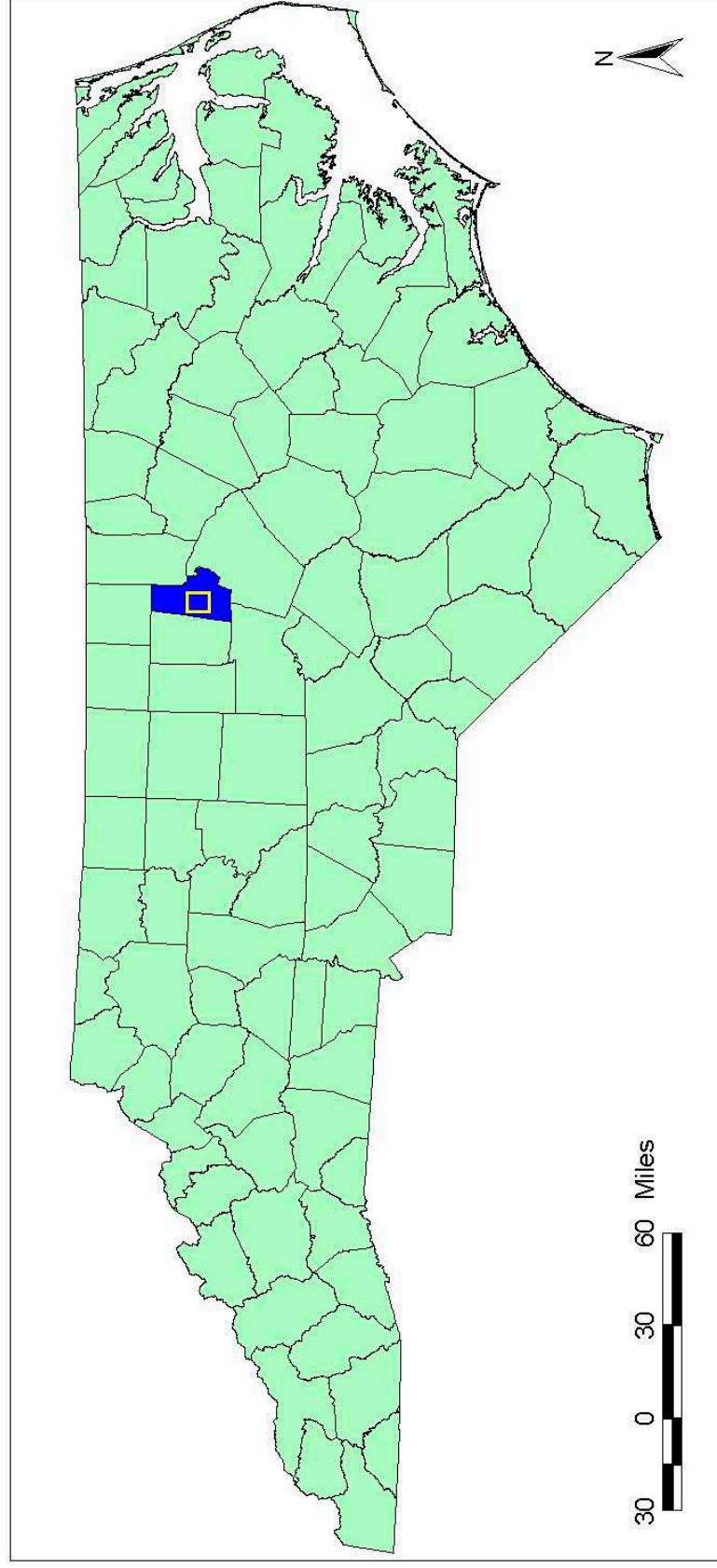


Figure 1

# GIS Modeling of Air Toxics Releases

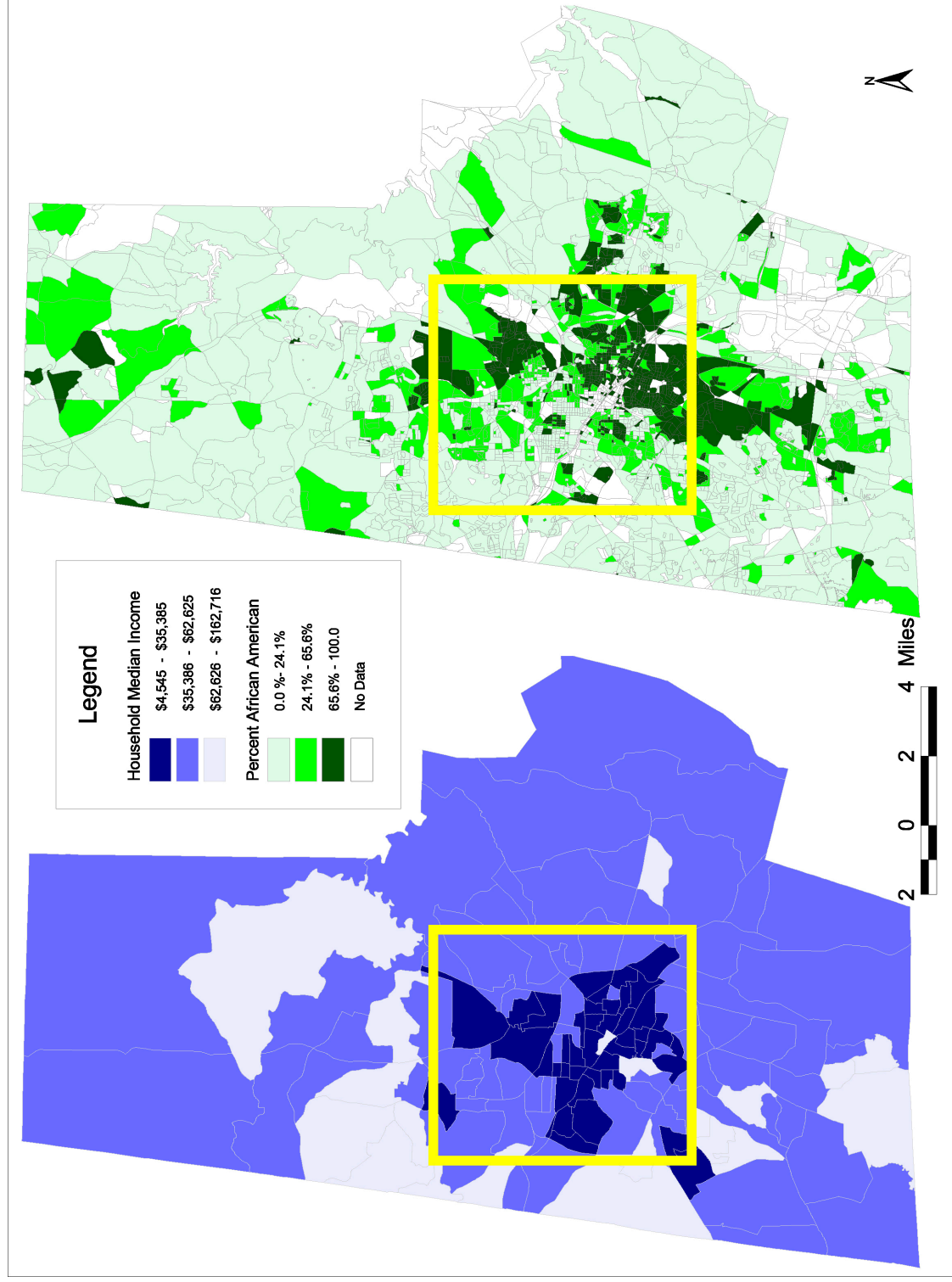


Figure 2

# GIS Modeling of Air Toxics Releases

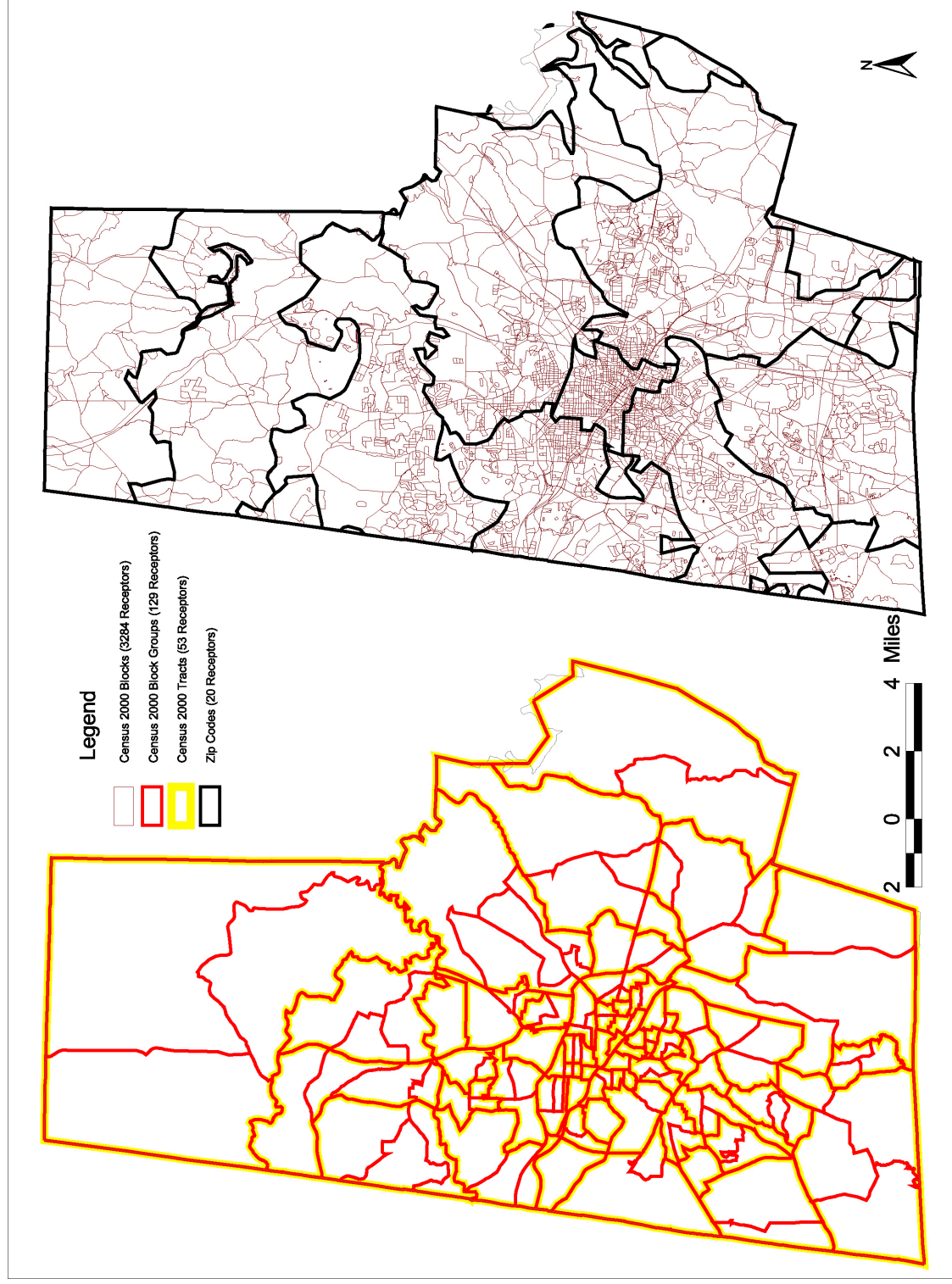


Figure 3

## GIS Modeling of Air Toxics Releases

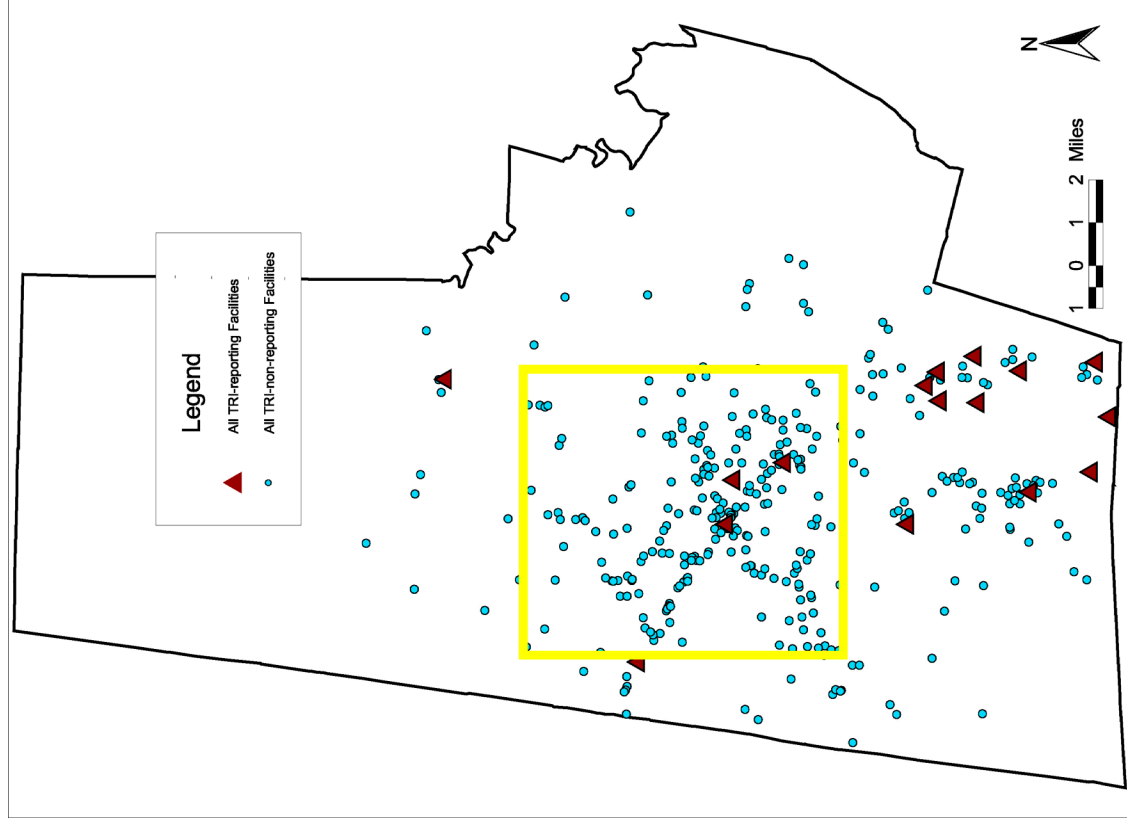


Figure 4

# GIS Modeling of Air Toxics Releases

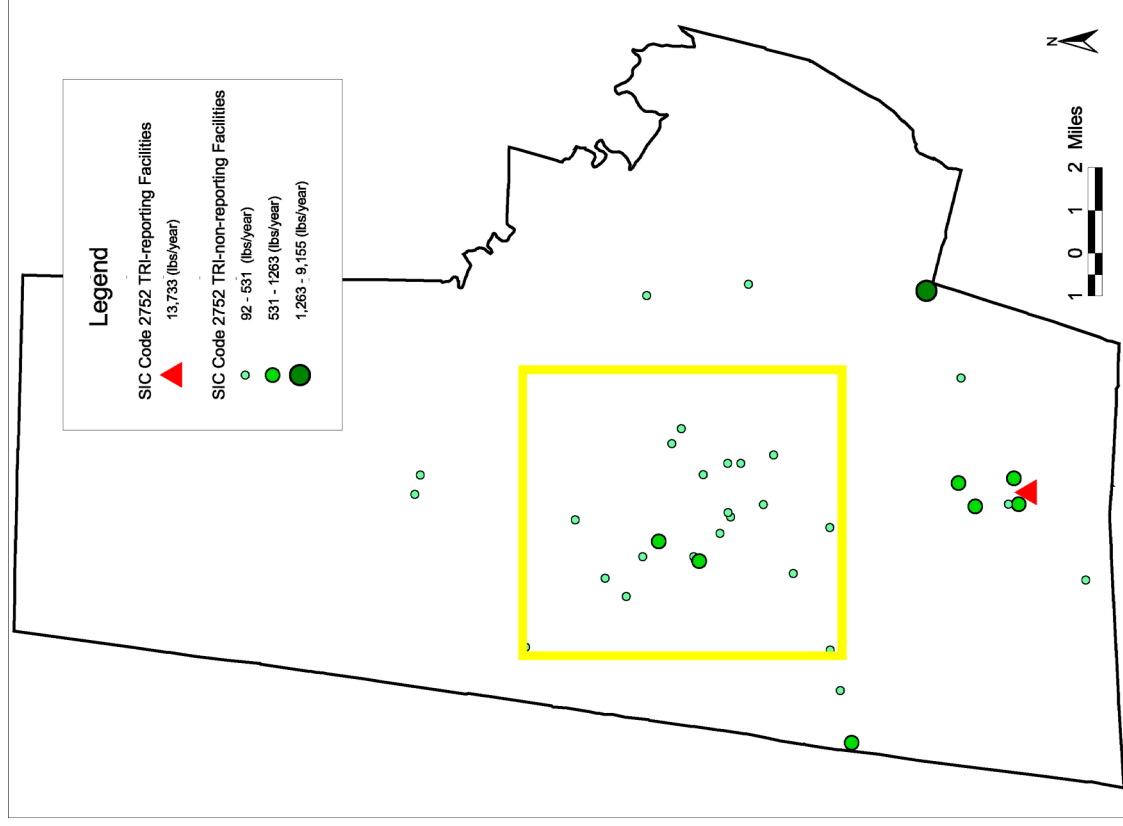


Figure 5

# GIS Modeling of Air Toxics Releases

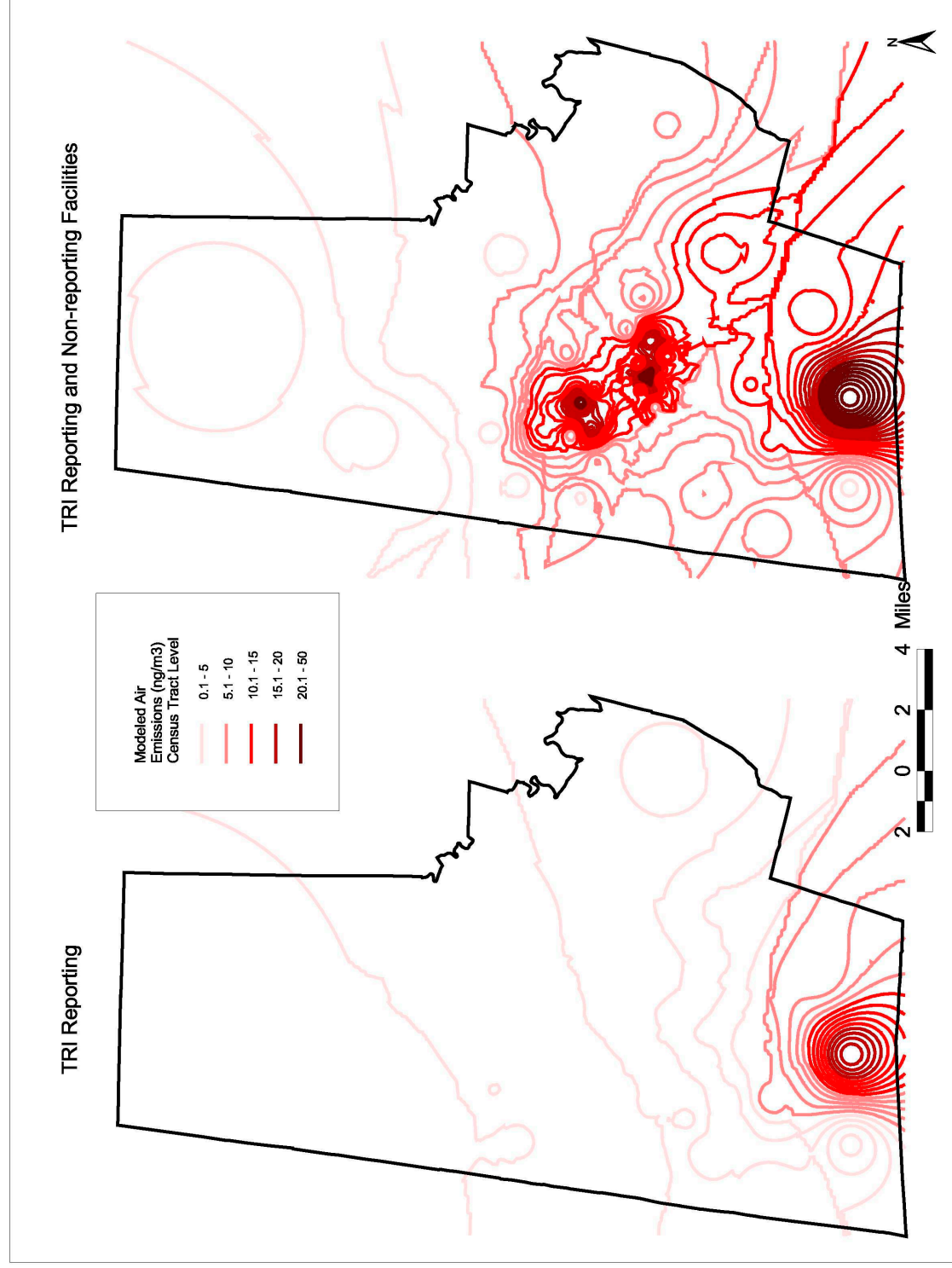


Figure 6

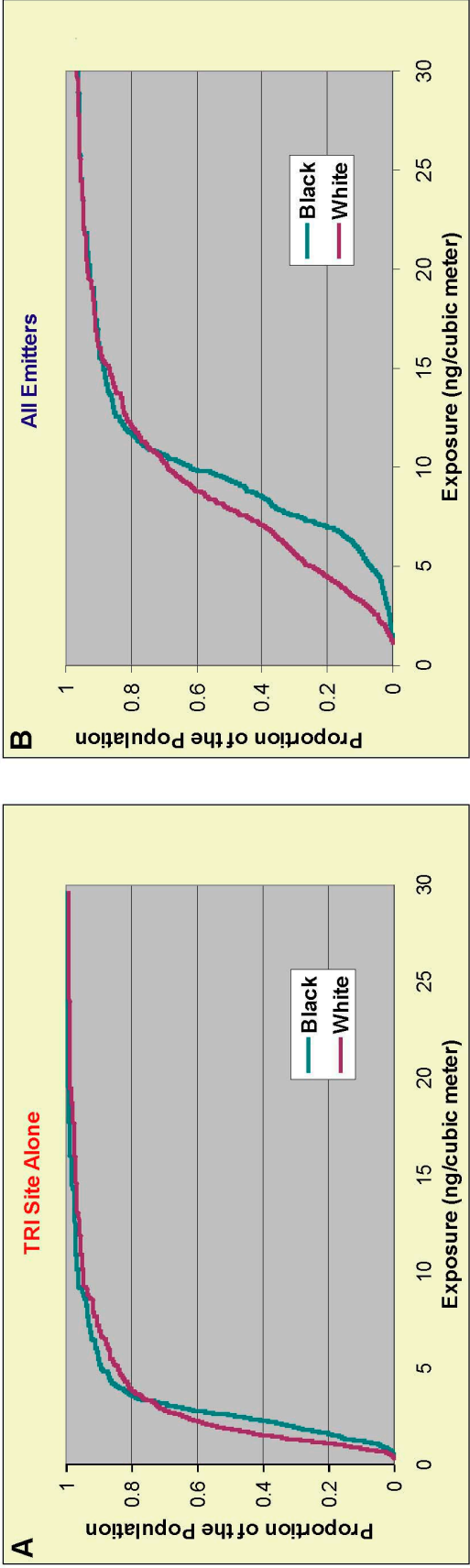


Figure 7



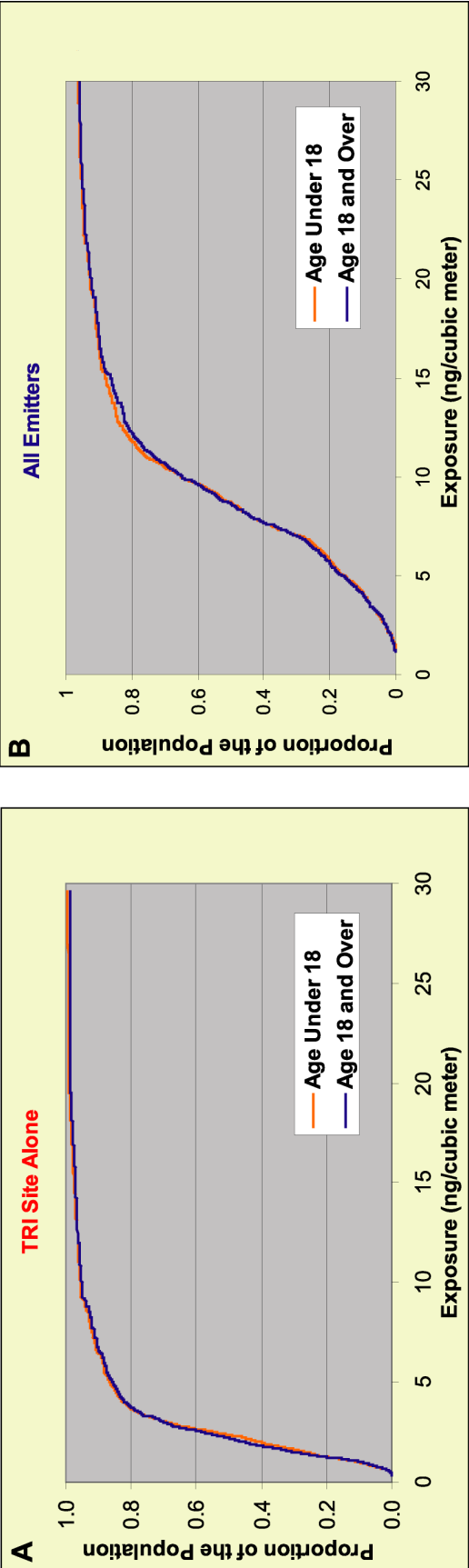


Figure 8



# GIS Modeling of Air Toxics Releases

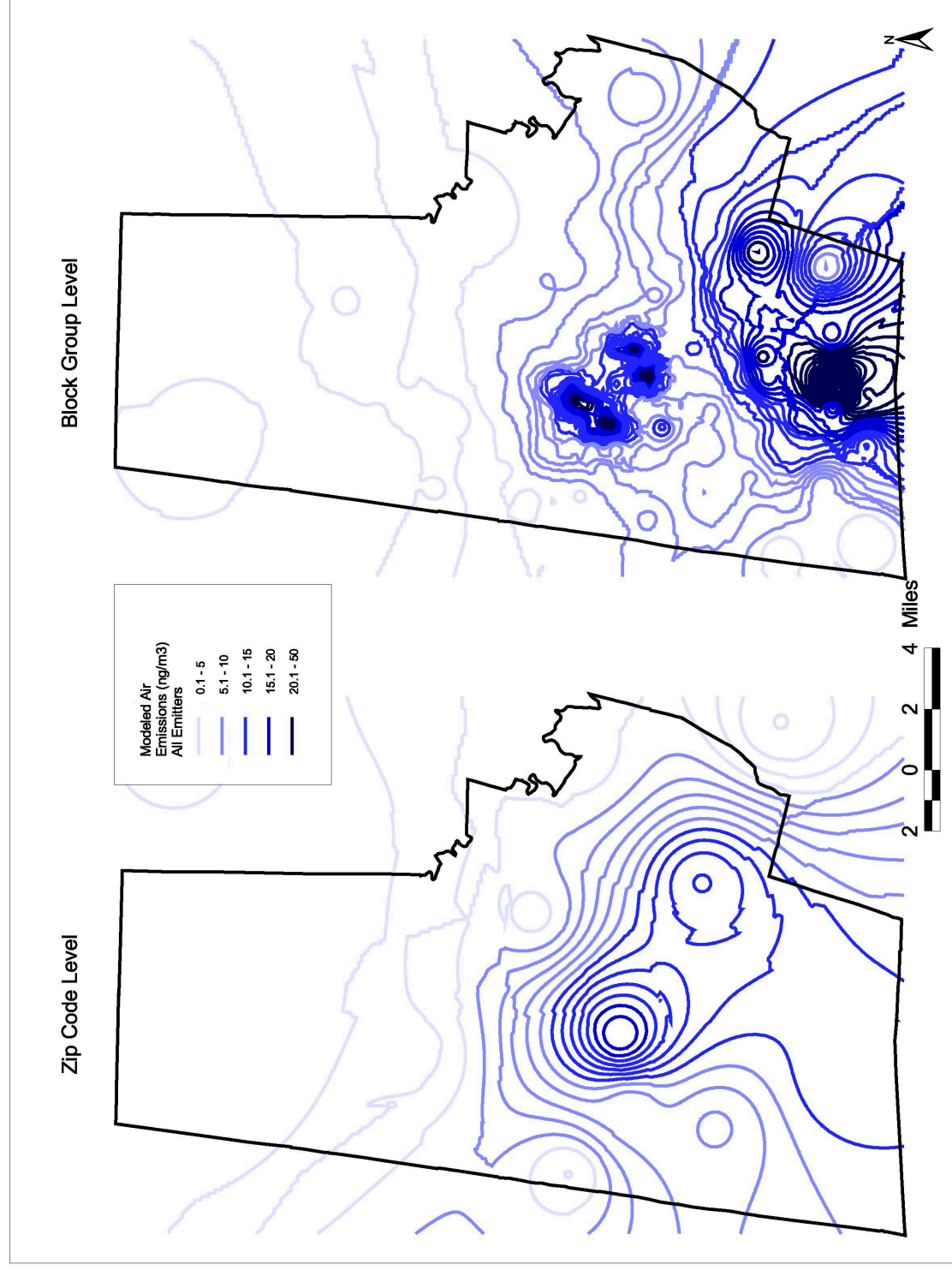


Figure 9